Proc. Natl. Acad. Sci. USA
Vol. 92, pp. 8308–8312, August 1995
Biophysics

Is water structure around hydrophobic groups clathrate-like?

TERESA HEAD-GORDON

Life Sciences Division, Lawrence Berkeley Laboratories, Berkeley, CA 94720

Communicated by Frank H. Stillinger, AT&T Bell Laboratories, Murray Hill, NJ, May 4, 1995 (received for review January 17, 1995)

ABSTRACT **The term "clathrate structure" is quantified for solvation of nonpolar groups by enumerating hydrogen-bonded ring sizes both in the solvation shell and through the shell–bulk interface and comparing it to a bulk control using the ST4 water model. For clathrate-like structure to be evident, the distributions along the hydrophobic surface are expected to be dominated by pentagons, with significant depletion of hexagons and larger polygons. While the distribution in this region is indeed distinguished by a large number of pentagons, there are significant contributions from hexagons and larger rings as well. Calculated polygon distributions through the shell–bulk interface indicate that when water structure is highly cooperative along the hydrophobic surface, hydrogen-bonded pathways leading back into bulk are then reduced. These results are qualitatively consistent with the observation that hydrophobicity is proportional to the nonpolar solute surface area.**

The connection between the thermodynamics of aqueous solvation of small hydrocarbons and the structure of water around these small entities is still not understood (1–12). Calorimetry experiments have shown that there is a proportionality between the unfavorable entropy signature at room temperature and hydrophobic surface area (1–4). Often "clathrate-like" behavior of water around these nonpolar molecules is invoked as a summarizing and simplified picture of the water structure attribute of this unfavorable entropy (1–4, 13–16), although the validity of such an analogy has not been characterized quantitatively.

The appropriateness of the term "clathrate behavior" as applied to water structure close to nonpolar guest molecules is analyzed directly. Non-short-circuited hydrogen-bonded pathways are calculated that close on themselves (17) for water molecules in the first solvation shell of the hydrophobic solutes, and the resulting polygon distribution is compared to that for arbitrary points in the liquid. There are striking differences between bulk liquid water—≈1:1 ratio of pentagons to hexagons, with evident contributions from both smaller and larger polygons (17)—and clathrates, where the pentagonal rings outnumber hexagons in a ratio of ≈8:1, with usually no lower and no higher order polygons present (18). This suggests that deviations toward larger numbers of pentagons and depletion of higher order polygons, especially hexagons, would be evident for water structure in the solvation shell surrounding nonpolar groups. For the process of hydrophobic association between two nonpolar solutes, further distinctions might be found defined by the non-short-circuited polygon counts for solute–solute separations corresponding to calculated free energy minima and barriers (7–12).

Models and Polygon Enumeration

A thermodynamic study of methane association at 298 K and 1.0 g/cm^3 by using both the ST4 model of water (19) and a new one-site potential for liquid water (12, 19, 20) have recently been presented. A detailed presentation of the potential models and simulation protocols have been discussed (12, 20); 200 coordinate sets were saved in the ST4 simulations for each perturbation window in order to generate the distributions discussed here. It is well established that rather small changes in the methane–water interaction parameters can result in reordering of the free energy minima where the solutes are either in contact or are so-called solvent separated (7, 8, 11). One aspect of the present study is to try to distinguish the structural origin, if any, of this reordering by comparing differences in polygon distributions generated from configurations using σ_{MeO} = 3.200 Å [favoring the solvent-separated region (7, 8)] and σ_{MeO} = 3.445 Å [favoring the contact minimum (7, 8, 11)].

In the region of the two methanes, a water molecule is considered to be in the first solvation shell if its oxygen center lies within 5.5 Å of either methane. The relative methane distances considered are 3.85, 5.25, and 6.875 Å, corresponding to regions near the contact, barrier, and solvent-separated stationary points on the free energy curve (7–12). When structurally distinguishing these regions from elsewhere in the solution, the following control is defined. An arbitrary point in the simulation box is chosen, and a second point is identified whose distance is equal to one of the above relative solute–solute distances. Around each of these two points, a sphere of radius 3.0 Å is drawn, and all water molecules whose oxygen centers lie within this region are eliminated. Any water molecule that lies within the spherical shell between 3.0 and 5.5 Å is designated a bulk control water. Several shells can be defined in this way, and 10 such control shells have been evaluated for all coordinate sets for the calculations reported here.

Non-short-circuited polygons are formed by starting on any molecule i and exhaustively searching for pathways defined by nearest neighbors, which after n steps return to molecule i (17). An energy criterion is used to define the nearest neighbors of molecule i—molecule j is a nearest neighbor if their interaction energy E_{ij} falls below some threshold V_c (17). The values V_c = −4.5, −4.0, −3.5, −3.0, and −2.5 kcal/mol (1 cal = 4.184 J) have been considered. Only the value V_c = −3.5 kcal/mol is reported since it provides the best balance between statistical significance and greatest structural differentiation. Given this value of V_c, a nearest-neighbor list for all water molecules in some chosen region can be made. Two regions have been considered to construct polygon distributions from a region along the hydrophobic surface and a region that interfaces with bulk. In the first region a nearest-neighbor list for a water molecule in the ≈2.5-Å shell constructed from all water molecules whose oxygen atoms are also present in the same shell and whose interaction energy falls below the threshold V_c is generated. This definition makes sense since the view that, compared to shells of water with no large excluded volume, clathrate cages might be present in the first solvation shell or that unusual structure might occur along the hydrophobic surface is being tested. The second region considers a nearest-neighbor list for all water molecules in the 2.5-Å shell constructed from all water molecules whose oxygen atoms are *not* present in the same shell, with an interaction energy below V_c. This region attempts to discriminate how the solvation shell

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

water interfaces with the bulk when unusual structure occurs along the hydrophobic surface.

Polygon counts are enumerated as follows. If a given water molecule has fewer than two nearest neighbors, it can never serve as a polygon vertex. Each molecule was individually considered and all polygons that terminate in three steps to form triangles were determined. All polygons that terminate in four steps were then determined for each molecule; whether these *ijkl* quadrilaterals are short-circuited by a nearest-neighbor *ik* interaction or *jl* interaction was carefully considered. If so, that quadrilateral is not counted, and the two resulting triangles have already been enumerated in the previous step. This procedure continues for successively larger polygons. The trends reported below do not change appreciably beyond 11-sided polygons, and I therefore have stopped polygon enumeration at this maximum size.

Results

The polygon distributions in Figs. 1–3 were generated from water configurations around two methanes separated by 3.875, 5.25, and 6.875 Å, respectively, using σ_{MeO} = 3.200 Å. Figs. 1*A*, 2*A*, and 3*A* compare polygon distributions with nonpolar groups (shaded bars) and without nonpolar groups (solid bars)

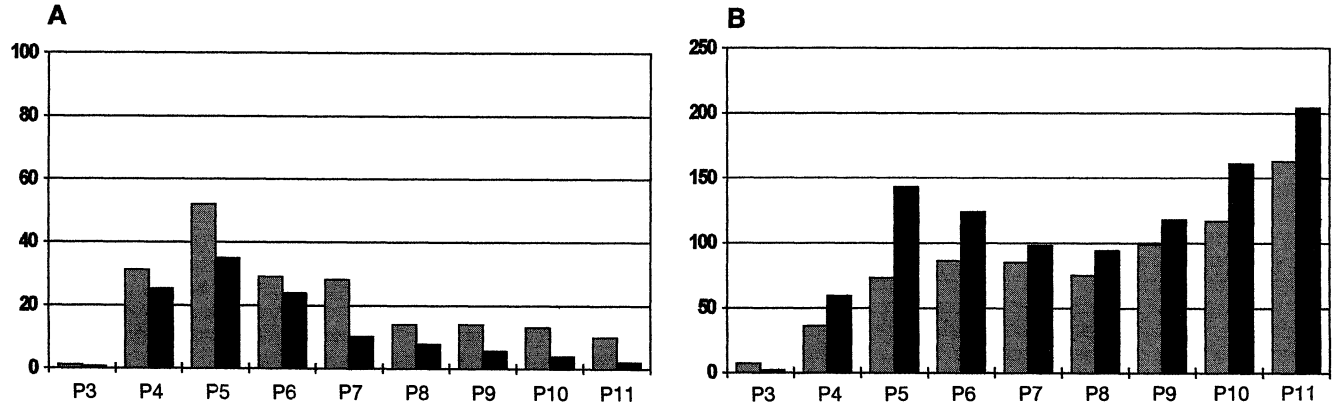

FIG. 1. Polygon enumeration for shell (shaded bars) and bulk control (solid bars) for free energy contact minimum using σ_{MeO} = 3.2 Å and ε_{MeO} = 0.2134 kcal/mol. (*A*) Polygon count along hydrophobic surface. (*B*) Polygon count through shell–bulk interface.

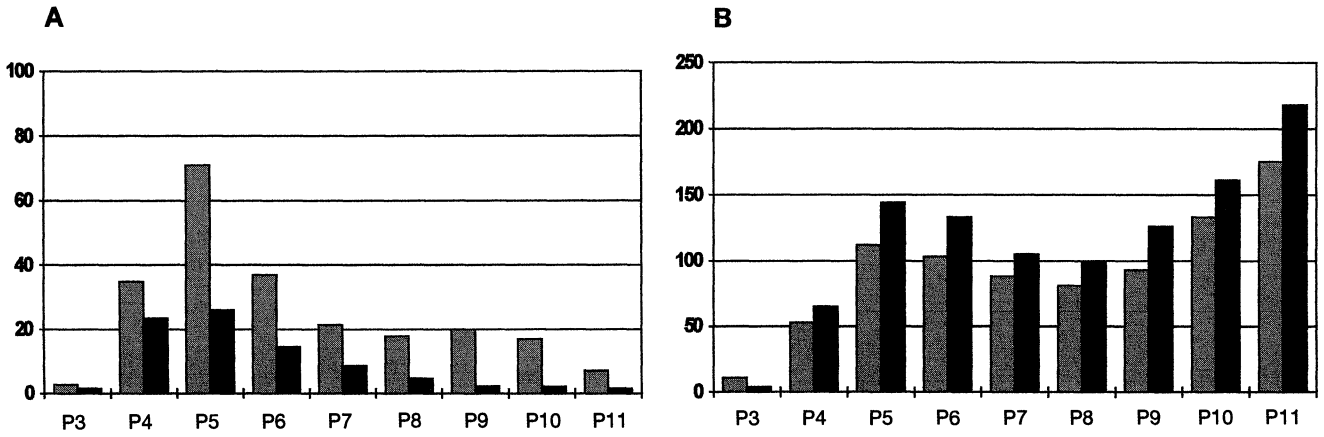

FIG. 2. Polygon enumeration for shell (shaded bars) and bulk control (solid bars) for free energy barrier using σ_{MeO} = 3.2 Å and ε_{MeO} = 0.2134 kcal/mol. (*A*) Polygon count for region along hydrophobic surface. (*B*) Polygon count through shell–bulk interface.

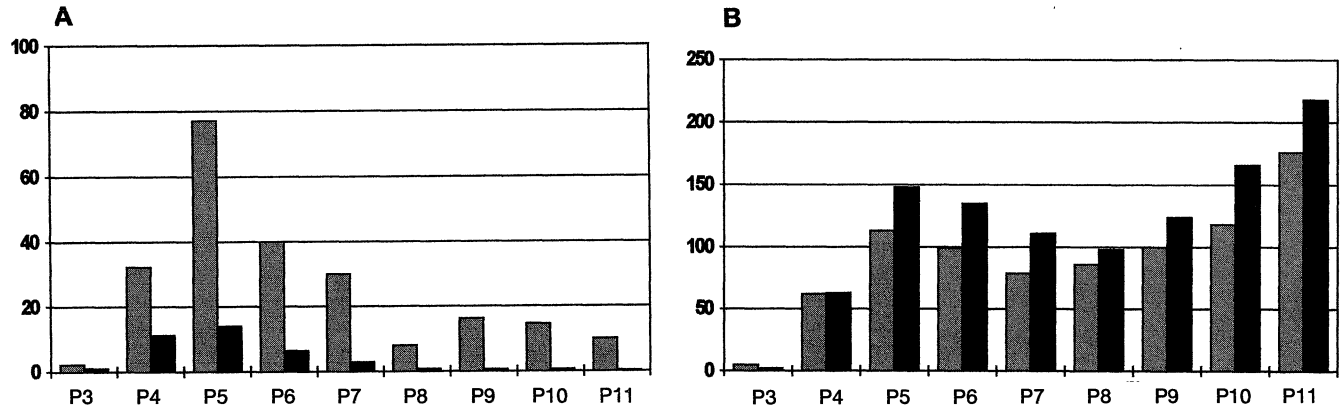

FIG. 3. Polygon enumeration for shell (shaded bars) and bulk control (solid bars) for free energy solvent-separated minimum using σ_{MeO} = 3.2 Å and ε_{MeO} = 0.2134 kcal/mol. (*A*) Polygon count for region along hydrophobic surface. (*B*) Polygon count through shell–bulk interface.

in the region along the hydrophobic surface. The most striking structural feature that emerges is the larger number of pentagons along the hydrophobic surface when hydrophobic groups are present, which is consistent with clathrate analogies. However, there are also significant contributions from hexagons and larger polygons (especially for larger methane–methane separations), which is inconsistent with clathrate behavior. Furthermore, the absolute numbers of all polygon sizes in a 2.5-Å shell along the hydrophobic surface are greater when hydrophobic groups are present, with the absolute number of pentagons being 1.5–5.0 times larger than the bulk control. By contrast, the bulk water control has networked within the shell quite differently. The number of tetramers is quite large relative to pentagons. This implies that without the presence of the hydrophobic solutes, the shell fulfills its higher order polygons through its interface with the bulk. Restricted to a definition of non-short-circuited polygons with all vertices present in the same shell, however, only short polygons are possible. The bulk control indicates that the water network is multidirectional around an arbitrary point, with directions penetrating the shell–bulk surface as important as networking within the shell.

This is indeed verified by comparing Figs. 1*A*, 2*A*, and 3*A* with Figs. 1*B*, 2*B*, and 3*B*. It is shown that the largest polygon (11-sided) dominates all distributions because of the definition of how nearest neighbors can be constructed through the shell–bulk interface. This definition does not permit other shell waters to be a nearest neighbor, which therefore eliminates much of the local structure that will give rise to smaller polygons. However, it is evident that fewer pentagons are present in the direction through the shell–bulk interface when hydrophobic groups are present compared to the bulk water control. Furthermore, the bulk control now shows the greater number of complete polygons compared to that when hydrophobic groups are present. Therefore, extensive networking within a water layer next to hydrophobic groups clearly restricts the number and type of hydrogen pathways that lead back into bulk water. The structural differences arising from the presence or absence of nonpolar groups is generic to all relative methane–methane distances and/or methane–water parameters. These results provide direct evidence that clathrate analogies are qualitatively valid in that there are larger numbers of pentagons. However, the analogy breaks down because of the still significant presence of larger polygons (especially hexagons), which also contribute to greater structural organization along the surface of excluded volume regions. It is also important to emphasize that the absolute number of any given polygon size is less than the number of coordinate sets used to generate the distributions, indicating that any organized structure is continuously formed and unformed.

A competition between exposed surface area and structural organization along the surface might give rise to free energy stationary points along a reaction coordinate of distance between two nonpolar solutes. The free energy minimum where the two solutes are in contact is entropically stabilized since fewer water molecules are involved compared to when they are separated by distances greater than ≈4.0 Å; this is true regardless of the methane–water parameter set used. Stationary points in the free energy at greater solute–solute separations would then arise from structural differences that destabilize or stabilize the barrier and the solvent-separated minimum where the solutes are separated by a water molecule or layer. It is evident from Figs. 1 and 3 that there is a significant increase in the number of pentagons in the favored solvent-separated region along the hydrophobic surface, and a corresponding increase in large polygons, compared to other regions on the reaction coordinate. Furthermore, the number of pathways describing pentagons leading back into bulk is most reduced in the contact region as well and is less diminished in the solvent-separated region. The polygon distributions generated in the barrier region show greater structural organization compared to the contact minimum. In fact, the distributions at the barrier more strongly resemble the solvent-separated region, although compared to that minimum the absolute numbers of pentagons, hexagons, and heptagons are fewer at the barrier. Apparently the competition between benefits gained by reducing surface area exposure and water organization along the surface is nonoptimal in either case at the barrier.

The polygon distributions in Figs. 4–6 are the same as in Figs. 1–3 except for the use of the methane–water parameter σ_{MeO} = 3.445 Å. The preferred contact minimum (Fig. 4) now shows comparable numbers of pentagons along the hydrophobic surface when compared to the solvent-separated region (Fig. 6). Furthermore, there is an increase in pentagons at the shell–bulk interface at the contact minimum (Fig. 4*B*) when compared to Fig. 1*B*. Apparently, the increase in clathrate-like structure using a larger interaction length parameter combined with reduced surface area now preferentially stabilizes the contact minimum. It is interesting to note that the polygon distributions generated in the barrier region (Fig. 5) now more strongly resemble the polygon counts of the contact minimum.

Conclusions

This work quantitatively characterizes the view that solvent ordering around hydrophobic solutes is due to clathrate behavior, by enumerating non-short-circuited polygons using the ST4 model, in the solvation shell around the hydrophobic groups compared to an appropriate bulk control. While pentagon contributions are much larger than the bulk control, quite significant contributions from larger rings also exist. It would seem that analogies with clathrates are useful by considering multidirectional sheets of water "pinched" by larger numbers of pentagons in the vicinity of the solutes to permit curling around them, analogous to the role that pentagons play in the structure of the fullerenes. The significant presence of polygon sizes larger than pentagons might also contribute to the ability of the network to enclose the solutes more effectively, since structural flexibility increases with ring size. The persistent presence of large polygons is obviously inconsistent with clathrate behavior and reveals that clathrate analogies are nonetheless too simple to explain the structural organization of water along hydrophobic surfaces. It has also been shown that greater structural organization along a hydrophobic surface will result in diminished numbers and types of pathways with which the shell waters can interface with the surrounding bulk.

It is worthwhile considering the thermodynamic consequences of the structural ordering in terms of closed hydrogen-bonded pathways reported here. (It is important to emphasize that no direct connection between structure and thermodynamics is made—i.e., through a formal statistical mechanical theory.) Because of structural organization along the surface, hydrogen-bonded pathways leading back into bulk clearly become more limited. This might suggest a molecular explanation for the observation that the hydrophobic effect, as measured by the large entropy change at room temperature, is proportional to the hydrophobic surface area, and the unfavorable entropy change may be partly due to restrictions of interfacing with bulk. An additional contribution to the unfavorable entropy of hydrophobic solvation is the greater rigidity of five-membered rings. Balancing this, entropic compensation is found by the greater flexibility of larger polygons—more so than a network of pentagons sharing sides to yield a rigid cage as in the case for clathrate hydrates.

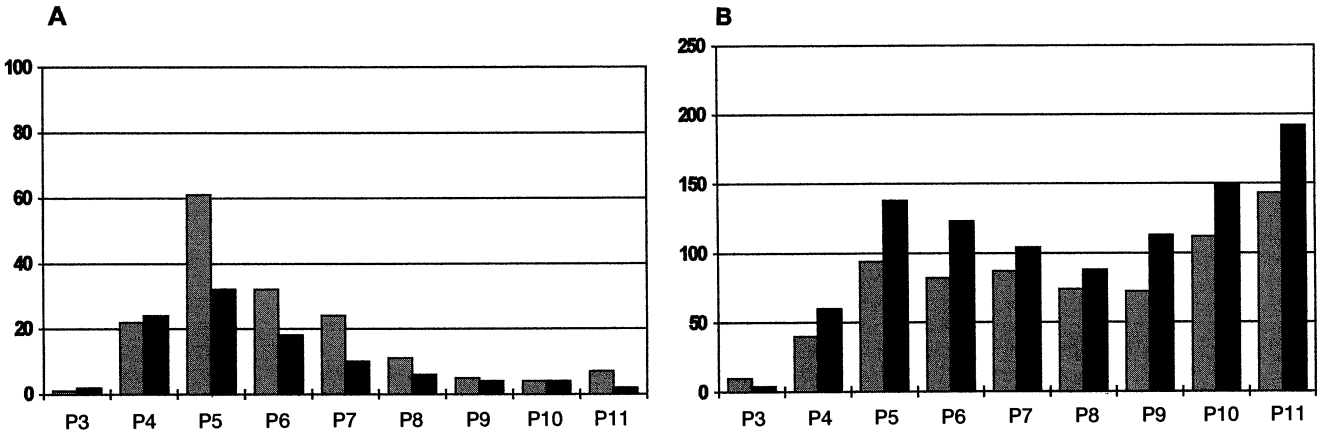

FIG. 4. Polygon enumeration for shell (shaded bars) and bulk control (solid bars) for free energy contact minimum using σ_{MeO} = 3.445 Å and ε_{MeO} = 0.2134 kcal/mol. (*A*) Polygon count for region along hydrophobic surface. (*B*) Polygon count through shell–bulk interface.

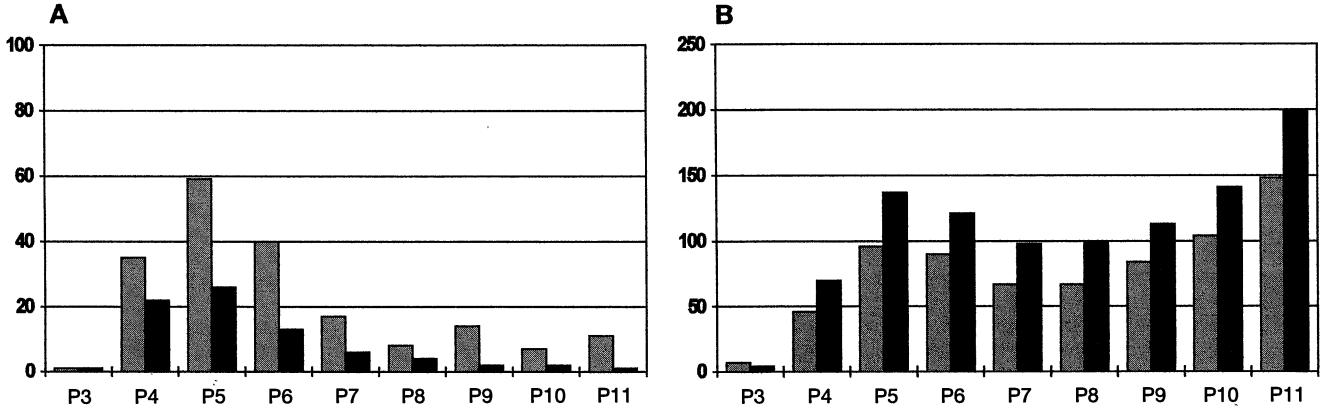

FIG. 5. Polygon enumeration for shell (shaded bars) and bulk control (solid bars) for free energy barrier using σ_{MeO} = 3.445 Å and ε_{MeO} = 0.2134 kcal/mol. (*A*) Polygon count for region along hydrophobic surface. (*B*) Polygon count through shell–bulk interface.

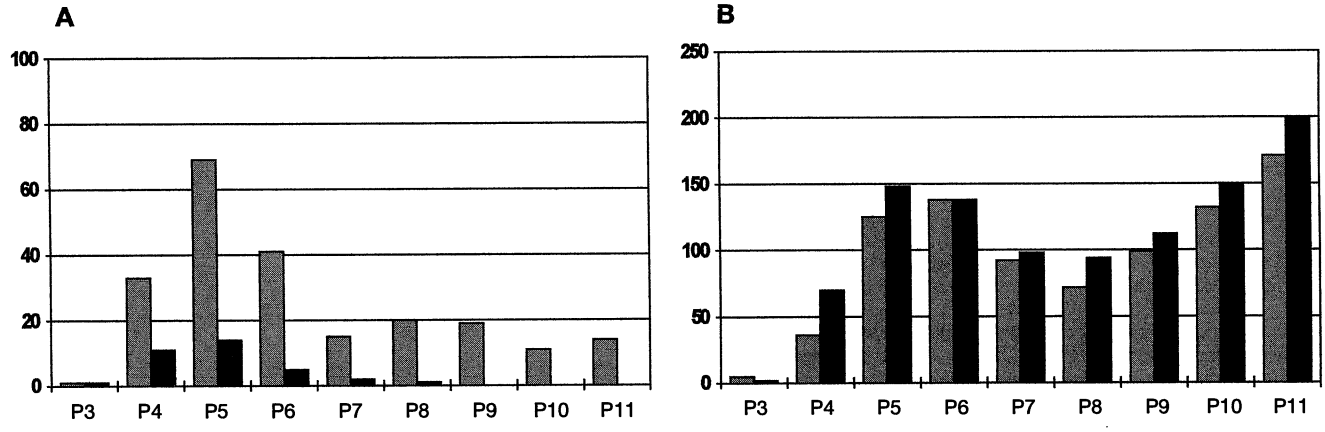

FIG. 6. Polygon enumeration for shell (shaded bars) and bulk control (solid bars) for free energy solvent-separated minimum using σ_{MeO} = 3.445 Å and ε_{MeO} = 0.2134 kcal/mol. (*A*) Polygon count for region along hydrophobic surface. (*B*) Polygon count through shell–bulk interface.

Calorimetry data on the transfer of hydrocarbons from their native state to water have clearly indicated that the hydrophobic effect is temperature dependent and that the entropy signature at room temperature gives way to enthalpy domination at higher temperatures (3, 4). Furthermore, the heat capacity change is large at room temperature, decreasing somewhat at higher temperatures (3, 4). The "melting" of ordered water structure in the vicinity of the nonpolar groups would certainly be consistent with the temperature dependence (greater orientational freedom at the expense of now broken hydrogen bonds) (1, 3). I might elaborate on this idea with more specifics—that hydrogen-bonded pathways involving large rings might be easier to maintain as temperature increases, while smaller five- and six-membered rings would be sacrificed. The loss of small rings would provide the largest gain in entropy and a definite decrease in enthalpy. Maintenance of only the larger rings would not be entropically costly at higher temperatures and would explain the decreased but still nonzero heat capacity at higher temperatures.

I thank Lawrence Pratt and Frank Stillinger for enjoyable discussions. This work is supported by the Office of Health and Environmental Research, Office of Energy Research, Department of Energy

(Grant DE-AC03-76-SF00098). A generous grant of Cray time from the National Energy Research Supercomputer Center is acknowledged.

1. Franks, F. (1973) in *Water: A Comprehensive Treatise*, ed. Franks, F. (Plenum, New York), Vol. 2, pp. 1–48.
2. Kauzmann, W. (1959) *Adv. Protein Chem.* **14,** 1–63.
3. Dill, K. A. (1990) *Biochemistry* **29,** 7133–7155, and references therein.
4. Privalov, P. L. & Gill, S. J. (1988) *Adv. Protein Chem.* **39,** 191–234.
5. Stillinger, F. H. (1980) *Science* **209,** 451–456.
6. Stillinger, F. H. (1973) *J. Solution Chem.* **2,** 141–158.
7. Pratt, L. R. & Chandler, D. (1977) *J. Chem. Phys.* **67,** 3683–3704.
8. Pratt, L. R. & Chandler, D. (1980) *J. Chem. Phys.* **73,** 3434–3441.
9. Pangali, C., Rao, M. & Berne, B. J. (1979) *J. Chem. Phys.* **71,** 2975–2981.
10. Jorgensen, W. L., Buckner, J. K., Boudon, S. & Tirado-Rives, J. (1988) *J. Chem. Phys.* **89,** 3742–3746.
11. Smith, D. E. & Haymet, A. D. J. (1993) *J. Chem. Phys.* **98,** 6445–6454.
12. Head-Gordon, T. (1995) *J. Am. Chem. Soc.* **117,** 501–507.
13. Pangali, C., Rao, M. & Berne, B. J. (1982) *J. Chem. Phys.* **81,** 2982–2990.
14. Zichi, D. A. & Rossky, P. J. (1985) *J. Chem. Phys.* **83,** 797–808.
15. Geiger, A., Rahman, A. & Stillinger, F. H. (1979) *J. Chem. Phys.* **70,** 263–279.
16. Ravishanker, G., Mezei, M. & Beveridge, D. L. (1982) *Faraday Symp. Chem. Soc.* **17,** 79–84.
17. Rahman, A. & Stillinger, F. H. (1973) *J. Am. Chem. Soc.* **95,** 7943–7949.
18. Davidson, D. W. (1973) in *Water: A Comprehensive Treatise*, ed. Franks, F. (Plenum, New York), Vol. 2, pp. 115–145.
19. Head-Gordon, T. (1994) *Chem. Phys. Lett.* **227,** 215–220.
20. Head-Gordon, T. & Stillinger, F. H. (1993) *J. Chem. Phys.* **98,** 3313–3327.